<u>**AMENDMENT**</u>

This listing of claims will replace all prior versions, and listings, of claims in the application:

<u>**Listing of Claims:**</u>

1-21. (Cancelled)

22. (Currently Amended) A ~~computer-implemented~~ method for a synthesis of photo-realistic animation of an object, the method comprising:

creating a first database of image samples showing the object in a plurality of appearances;

creating a second database of [[the]] visual features for each image sample of the object;

creating a third database of [[the]] non-visual features of the object in each image sample;

obtaining, for each frame in a plurality of N frames of an object animation, a target feature vector comprising visual features and non-visual features associated with the object animation;

selecting, via a processor <u>and an audio/video unit selection process, longest possible</u> candidate image samples from the first database utilizing the target feature vector to generate the photo-realistic animation of the object;

calculating, for each pair of candidates of two consecutive frames, a concatenation cost from a combination of <u>the</u> visual features from the second database and <u>the non-visual features</u> from the third database; and

performing a Viterbi search to find a least expensive path through the <u>each pair of</u> candidates accumulating a target cost and <u>the</u> concatenation <u>cost</u> ~~costs, wherein generating the~~

~~photo-realistic animation of the object occurs using an audio/video unit selection process in which a longest possible candidate image sample is selected~~.

23. (Currently Amended) A ~~computer-implemented~~ method for the synthesis of photo-realistic animation of an object, the method utilizing a first database of image samples showing an object in a plurality of appearances, a second database of visual features for each image sample of the object, and a third database of non-visual features of the object ~~in each image sample~~, the method comprising:

obtaining, for each frame in a plurality of N frames of an object animation, a target feature vector comprising visual features and non-visual features associated with the object animation;

selecting, via a processor <u>and an audio-video unit selection process, longest possible</u> candidate image samples from the first database utilizing the target feature vector to generate a photo-realistic animation of the object;

calculating, for each pair of candidates of two consecutive frames, a concatenation cost from a combination of <u>the</u> visual features from the second database and ~~object characteristics~~ <u>the non-visual features</u> from the third database; and

performing a Viterbi search to find [[the]] <u>a</u> least expensive path through the <u>each pair of</u> candidates accumulating a target cost and <u>the</u> concatenation <u>cost</u> ~~costs, wherein generating the photo-realistic animation of the object occurs using an audio/video unit selection process in which a longest possible candidate image sample is selected~~.

24. (Currently Amended) The method of claim 22, wherein selecting candidate image samples further comprises for each frame in the plurality of N frames of the animation, selecting

candidate image samples associated with the object animation using a comparison of a combination of the visual features and the non-visual features with the target feature vector.

25. (Currently Amended) The method of claim 24, further comprising compiling the ~~selected~~ candidate image ~~sample candidates~~ samples to form a photo-realistic animation.

26-28. (Cancelled)

29. (Previously Presented) The method of claim 22, wherein the animation is a talking-head animation, the first database stores sample images of a face that speaks, the second database stores associated facial visual features and the third database stores acoustic information for each frame in the form of phonemes.

30. (Currently Amended) ~~A computer-implemented~~ A non-transitory computer-readable media storing an animation of an object generated according to a process of:

creating a first database of image samples showing an object in a plurality of appearances;

creating a second database of [[the]] visual features for each image sample of the object;

creating a third database of [[the]] non-visual features of the object in each image sample;

obtaining, for each frame in a plurality of N frames of an object animation, a target feature vector comprising visual features and non-visual features associated with the object animation;

selecting, via a processor and an audio/video unit selection process, longest possible

candidate image samples from the first database utilizing the target feature vector to generate a

photo-realistic animation of the object;

calculating, for each pair of candidates of two consecutive frames, a concatenation cost

from a combination of ~~the~~ visual features from the second database and ~~object characteristics~~ the

non-visual features from the third database; and

performing a Viterbi search to find the least expensive path through the candidates

accumulating a target cost and ~~the~~ concatenation cost ~~costs, wherein generating the photo-

realistic animation of the object occurs using an audio/video unit selection process in which a

longest possible candidate image sample is selected~~.

31. (Currently Amended) The ~~animation~~ non-transitory computer-readable medium of claim 30,

wherein the step of selecting candidate image samples further comprises for each frame in the

plurality of N frames of the animation, selecting candidate image samples associated with the

object animation using a comparison of a combination of ~~the~~ visual features and ~~the~~ non-visual

features with the target feature vector.

32. (Currently Amended) The ~~animation~~ non-transitory computer-readable medium of claim 31,

wherein the animation is generated according to the additional step of compiling the ~~selected~~

candidate image ~~sample candidates~~ samples to form a photo-realistic animation.

33-35. (Cancelled)

36. (Currently Amended) A system for synthesizing a photo-realistic animation of an object, the

system comprising:

a processor;

a first module controlling the processor to create a first database of image samples showing the object in a plurality of appearances;

a second module controlling the processor to create a second database of [[the]] visual features for each image sample of the object;

a third module controlling the processor to create a third database of [[the]] non-visual features of the object in each image sample;

a fourth module controlling the processor to obtain, for each frame in a plurality of N frames of an object animation, a target feature vector comprising visual features and non-visual features associated with the object animation;

a fifth module controlling the processor to select, via a processor and an audio/video unit selection process, longest possible candidate image samples from the first database utilizing the target feature vector to generate the photo-realistic animation of the object;

a sixth module controlling the processor to calculate, for each pair of candidates of two consecutive frames, a concatenation cost from a combination of the visual features from the second database and object characteristics the non-visual features from the third database; and

a seventh module controlling the processor to perform a Viterbi search to find a least expensive path through the candidates accumulating a target cost and the concatenation cost costs, wherein generating the photo-realistic animation of the object occurs using an audio/video unit selection process in which a longest possible candidate image sample is selected.

37. (Currently Amended) A non-transitory computer-readable storage medium storing instructions which, when executed by a computing device, cause the computing device to synthesize photo-realistic animation of an object, the instructions utilizing a first database of

image samples showing an object in a plurality of appearances, a second database of visual features for each image sample of the object, and a third database of non-visual features of the object ~~in each image sample~~, the instructions comprising:

obtaining, for each frame in a plurality of N frames of an object animation, a target feature vector comprising visual features and non-visual features associated with the object animation;

selecting, via a processor <u>and an audio/video unit selection process, longest possible</u> candidate image samples from the first database utilizing the target feature vector to generate a photo-realistic animation of the object;

calculating, for each pair of candidates of two consecutive frames, a concatenation cost from a combination of <u>the</u> visual features from the second database and ~~object characteristics~~ <u>the</u> <u>non-visual features</u> from the third database; and

performing a Viterbi search to find the least expensive path through the candidates accumulating a target cost and <u>the</u> concatenation <u>cost</u> ~~costs, wherein generating the photo-realistic animation of the object occurs using an audio/video unit selection process in which a longest possible candidate image sample is selected~~.

38. (Currently Amended) The non-transitory computer-readable storage medium of claim 37, wherein selecting candidate image samples further comprises for each frame in the plurality of N frames of the animation, selecting candidate image samples associated with the object animation using a comparison of a combination of <u>the</u> visual features and <u>the</u> non-visual features with the target feature vector.

39. (Currently Amended) The non-transitory computer-readable storage medium of claim 38, further comprising compiling the ~~selected~~ candidate image ~~sample candidates~~ samples to form a photo-realistic animation.

40. (Previously Presented)　　The non-transitory computer-readable storage medium of claim 37, wherein the animation is a talking-head animation, the first database stores sample images of a face that speaks, the second database stores associated facial visual features and the third database stores acoustic information for each frame in the form of phonemes.

41. (Currently Amended) A system for animating an object, the system utilizing a first database of image samples showing an object in a plurality of appearances, a second database of visual features for each image sample of the object, and a third database of non-visual features of the object ~~in each image sample~~, system comprising:

　　　　a processor; and

　　　　a first module controlling the processor to animate an object generated according to steps comprising:

　　　　　　　　obtaining, for each frame in a plurality of N frames of an object animation, a target feature vector comprising visual features and non-visual features associated with the object animation;

　　　　　　　　selecting, via the processor and using an audio/video unit selection process, longest possible candidate image samples from the first database utilizing the target feature vector to generate a photo-realistic animation of the object;

　　　　　　　　calculating, for each pair of candidates of two consecutive frames, a concatenation cost from a combination of the visual features from the second database and ~~object~~

~~characteristics~~ the non-visual features from the third database; and

performing a Viterbi search to find the least expensive path through the candidates accumulating a target cost and the concatenation cost ~~costs, wherein generating the photo-realistic animation of the object occurs using an audio/video unit selection process in which a longest possible candidate image sample is selected~~.

42. (Currently Amended) The system of claim 41, wherein selecting candidate image samples further comprises for each frame in the plurality of N frames of the animation, selecting candidate image samples associated with the object animation using a comparison of a combination of the visual features and the non-visual features with the target feature vector.

43. (Currently Amended) The system of claim 42, wherein the animation is generated according to the additional step of compiling the ~~selected~~ candidate image ~~sample candidates~~ samples to form a photo-realistic animation.